

Supplementary Information:
Statistical Mechanics of Dimerizations
and its Consequences for Small Systems

Ronen Zangi*^{1,2}

¹*POLYMAT & Department of Organic Chemistry I, University of the Basque Country UPV/EHU,
Avenida de Tolosa 72, 20018, Donostia-San Sebastián, Spain*

²*IKERBASQUE, Basque Foundation for Science, Plaza Euskadi 5, 48009 Bilbao, Spain*

November 2, 2022

SI-1 Analytical Evaluations of K

The equilibrium constant of dimerization derived in the main text and expressed in Eq. 20 in terms of ensemble average of reactant and product concentrations is now compared against two analytical evaluations based on the single-particle, q_A , and pair-particle, q_{A_2} , partition functions. To this end, we simplify our system and model the reactants, A , only as single-site particles, thus, removing the protecting site that prevented higher-order aggregation. To preclude the formation of aggregates larger than a dimer, we simply restrict this test system to $N_A^{\text{total}} = 2$. We choose the a cubic box with $L_{\text{box}} = 6.0 \text{ nm}$ thus $c_A^{\text{total}} = 0.00926 \text{ molecule/nm}^3$. To render the magnitude, as well as the location, of the first maximum of $g(r)$ in the single-site system and in the main-model system similar, we modified ϵ and σ parameters of the LJ potential to $\epsilon^{LJ} = 26.90 \text{ kJ/mol}$ and $\sigma = 0.152 \text{ nm}$. Other simulation parameters were unchanged. The MC simulation consisted of $8 \cdot 10^{12}$ trial moves whereas the MD simulation was run for $720 \mu\text{s}$. The value of K obtained by Eq. 20, for each of these simulations, is listed in Table SI-1.2.

I. K from Integration over Particle's Coordinates

In this approach we completely separate the integrations over momenta from those over spatial coordinates. If \mathcal{T} is the kinetic part of the Hamiltonian, the single-particle partition function of unbound A can be written as,

$$q_A(r) = \frac{1}{h^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\beta\mathcal{T}(\vec{p}_A)} d\vec{p}_A \int_{r_A} d\vec{r}_A = \frac{V}{h^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\beta\mathcal{T}(\vec{p}_A)} d\vec{p}_A \quad , \quad (\text{SI-1.1})$$

where h is Planck's constant and the integral over r_A is of three dimensions yielding the volume when the particle does not interact with its surrounding. If \mathcal{U} is the potential part of the Hamiltonian and r_c the cutoff distance defining the bound state, the pair-particle partition function can be written as,

$$q_{A_2}(\vec{p}_{A'}, \vec{p}_{A''}, \vec{r}_{A'}, \vec{r}_{A''}) = \frac{1}{h^6} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\beta\mathcal{T}(\vec{p}_{A'}, \vec{p}_{A''})} d\vec{p}_{A'} d\vec{p}_{A''} \int_{r_{A'}} d\vec{r}_{A'} \int_0^{r_c} e^{-\beta\mathcal{U}(r)} 2\pi r^2 dr \quad , \quad (\text{SI-1.2})$$

where we labeled the first particle A' , and the second A'' . The relative distance is defined as $r = |\vec{r}_{A''} - \vec{r}_{A'}|$, and the usual volume element for integration over this relative distance, $4\pi r^2 dr$,

is divided by two because A' and A'' are indistinguishable. In addition, the spatial integration over the coordinates of the first particle A' gives V , thus, the equilibrium constant can be expressed as,

$$K = \frac{q_{A_2}}{q_A^2} V c^\emptyset = \frac{1}{2} c^\emptyset \int_0^{r_c} e^{-\beta u(r)} d\vec{r} = c^\emptyset \int_0^{r_c} e^{-\beta u(r)} 2\pi r^2 dr \quad , \quad (\text{SI-1.3})$$

where the integrals over momenta cancel-out when taking the ratio of the partition functions. Equation SI-1.3 can be solved numerically and the result is shown in Table SI-1.2.

II. K from a Molecular Partition Function

We now evaluate q_{A_2} by integrations over coordinates and momenta of the center-of-mass of the dimer and over the relative motions therein. This is realized by writing the Hamiltonian of the pair-particle partition function in terms of generalized coordinates and momenta that describe translation of center-of-mass, as well as, rotation and vibration of the bound state. If the rotational and vibrational modes are decoupled, the expression of K becomes¹,

$$K = \frac{q_{\text{trans}}(A_2) \cdot q_{\text{rot}}(A_2) \cdot q_{\text{vib}}(A_2) \cdot e^{-\beta \epsilon}}{q_{\text{trans}}^2(A)} V c^\emptyset \quad , \quad (\text{SI-1.4})$$

where ϵ equals $-\epsilon^{LJ}/N_{\text{Avogadro}}$. In the 'classical' approximation, where the sum over translational states can be substituted by an integral, the translational partition function has the form,

$$q_{\text{trans}}^{\text{classical}} = \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} V \quad , \quad (\text{SI-1.5})$$

where m is the mass of the moving body. The rotational partition function of a homonuclear rigid-rotor dimer at high-temperatures is,

$$q_{\text{rot}} = \frac{8\pi^2 I k_B T}{2h^2} \quad , \quad (\text{SI-1.6})$$

where the moment of inertia is $I = \mu R_{eq}^2$, μ the reduced mass, and $R_{eq} = 0.1707 \text{ nm}$ the equilibrium bond length of the dimer. The evaluation of the vibrational partition function is normally preceded by an input of the vibrational frequency (or force-constant). Because the vibrations in our dimer are actually oscillatory motions around the minimum of the LJ potential, we also apply the high-temperature approximation in this case and evaluate the vibrational partition function by performing numerical integration instead of discrete summation. The Hamiltonian here includes a

one-dimensional kinetic term of a body with a reduced mass μ and the LJ potential is shifted by ϵ^{LJ} so its minimum is at zero energy. Consequently we get,

$$q_{\text{vib}} = \frac{1}{h} \int_{-\infty}^{\infty} e^{-\beta p^2/2\mu} d\vec{p} \int_0^{r_c} e^{-\beta[U_{LJ}(r)+\epsilon^{LJ}]} dr = \left(\frac{2\pi\mu k_B T}{h^2} \right)^{1/2} \int_0^{r_c} e^{-\beta[U_{LJ}(r)+\epsilon^{LJ}]} dr, \quad (\text{SI-1.7})$$

which can be easily calculated. The values of the different terms of the molecular partition function of the dimer are exhibited in Table SI-1.1.

Table SI-1.1: The values of different modes in the molecular partition function of the dimer, along with the corresponding monoatomic partition function and the Boltzmann's factor, necessary to compute the equilibrium constant of our test system ($V = 216.0 \text{ nm}^3$ and $T = 300.0 \text{ K}$) via Eq. SI-1.4.

$q_{\text{trans}}^{\text{classical}}(A_2)$	q_{rot}	q_{vib}	$e^{-\beta\epsilon}$	$q_{\text{trans}}^{\text{classical}}(A)$
$1.8866 \cdot 10^7$	90.103	0.36473	48261	$6.6702 \cdot 10^6$

The comparisons between the equilibrium constant, as well as of the standard Gibbs energy change, obtained by the four different (two simulation- and two analytical-) methods is shown in Table SI-1.2. The agreement between the MC simulation and the numerical integration over particles' coordinates (Eq. SI-1.3) is almost perfect. Relative to this, the agreement of K between the MC and MD simulations may seem compromised. However when considering the difference between the corresponding ΔG^\varnothing , which equals 0.02 kJ/mol , the agreement is still very good, and the mild discrepancy can be attributed to application of a thermostat to a system with small number of degrees of freedom. By far, the largest deviation is observed when the calculation is performed using the molecular partition function (Eq. SI-1.4) where the difference in ΔG^\varnothing with the other methods is in the range $0.06 - 0.09 \text{ kJ/mol}$. As we argued before², this is not surprising given the several assumptions made in deriving this equation, and in particular, the neglect of coupling between vibrational and rotational modes for a bond formed by a 'soft', intermolecular, potential.

Table SI-1.2: Comparison between values of the equilibrium constant K computed by four different methods, for the dimerization described in Eq. 2 of the simplified model system of single-site monomers detailed in this section. In the two simulation methods, Monte-Carlo (MC) and Molecular Dynamics (MD), K was obtained by calculating the ratio between the product and correlated-reactants concentration according to Eq. 20. The analytical/numerical calculations were based on integration of the particles coordinates (Eq. SI-1.3), as well as on partition functions describing relative motions of a homonuclear diatomic molecule (Eq. SI-1.4). In addition to the value of K , we also display (in kJ/mol) the corresponding change in the standard Gibbs energy, ΔG° , using the definition in Eq. 5.

	Simulations (Eq. 20)		Analytical/Numerical Calculations	
	MC	MD	Eq. SI-1.3	Eq. SI-1.4
K	90.625 \pm 0.005	89.73 \pm 0.24	90.623	87.481
ΔG°	-11.2413 \pm 0.0001	-11.217 \pm 0.007	-11.2412	-11.1532

SI-2 Limits on the Relation between Reference and Finite Systems

The relation expressed in Eq. 12 between partition functions of the reference state and those of the arbitrary system assumes translational partition functions of monomer and dimer are linearly proportional to the volume. This is true if these translational partition functions can be described 'classically' as considered in Eq. SI-1.5. For macroscopic reference systems this assumption is clearly valid. However, would it also hold for a chosen system that is finite in size, thus, with a small volume?

In obtaining Eq. SI-1.5, quantum translational energy states are actually considered however the discrete sum, that in 1-dimension (along the x -axis) takes the form¹

$$q_{\text{trans},x} = \sum_{n_x=1}^{\infty} \exp \left[-\beta h^2 n_x^2 / (8mL_{\text{box}}^2) \right] \quad (\text{SI-2.1})$$

with n_x a positive integer, is approximated by an integral over n_x ,

$$q_{\text{trans},x} \approx \int_0^{\infty} \exp \left[-\beta h^2 n_x^2 / (8mL_{\text{box}}^2) \right] dn_x \quad . \quad (\text{SI-2.2})$$

Because motion along each axes is independent, the translational partition function in 3-dimensions becomes,

$$q_{\text{trans}} = q_{\text{trans},x} \cdot q_{\text{trans},y} \cdot q_{\text{trans},z} \quad . \quad (\text{SI-2.3})$$

Approximating Eq. SI-2.1 by Eq. SI-2.2 requires successive terms in the sum to be spaced close enough. In fact, the spacing is constant with a value of an integer unit, nonetheless, it can be small *relative* to the range (width along the n_x axis) of significant terms that are summed. Given the Gaussian form of the terms inside the sum, the condition is that the width $\sigma = \sqrt{(8mL_{\text{box}}^2)/(\beta h^2)}$ should be much larger than 1. For the single-site monomer system mentioned in Section SI-1 ($m = 10 \text{ amu}$, $L_{\text{box}} = 6.0 \text{ nm}$, and $T = 300 \text{ K}$), the value of σ is 212. Although this may be considered a large number compared to 1, we also assess the approximation directly by calculating q_{trans} (Eq. SI-2.3) using the discrete summation of energies as indicated in Eq. SI-2.1. The results are, $q_{\text{trans}(A)} = 6.6171 \cdot 10^6$ and $q_{\text{trans}(A_2)} = 1.8760 \cdot 10^7$, for the monomer and dimer respectively. The corresponding values using the 'classical' translation approximation (Eq. SI-1.5), shown in Table SI-1.1, exhibit relative deviations of 0.6 % and 0.8 %. As a matter of fact, our aim is to

assess the 'classical' approximation applied to the ratio of the partition functions shown in Eq. 12.

We therefore define,

$$R_q = \frac{q_{\text{trans}}(A_2)}{[q_{\text{trans}}(A)]^2}, \quad (\text{SI-2.4})$$

as well as the corresponding ratio of the 'classical' translational partition functions,

$$R_q^{\text{classical}} = \frac{q_{\text{trans}}^{\text{classical}}(A_2)}{[q_{\text{trans}}^{\text{classical}}(A)]^2}, \quad (\text{SI-2.5})$$

and quantify the relative error by,

$$\Delta = \frac{R_q - R_q^{\text{classical}}}{R_q}, \quad (\text{SI-2.6})$$

which also represents the relative error in determining K . This gives $\Delta = 0.010$, thus an error of 1 %, an acceptable accuracy for many applications.

As apparent from Eq. SI-2.1, besides volume, the width of the translational energy sum is also affected by mass and temperature. To study the effect of these three parameters systematically, we consider the single-site monomer system again and vary each parameter while keeping the other two constants. We then plot Δ as a function of the parameter that is changed and display the results in Fig. SI-2.1. As expected Δ decreases for heavier masses, higher temperatures, and larger volumes.

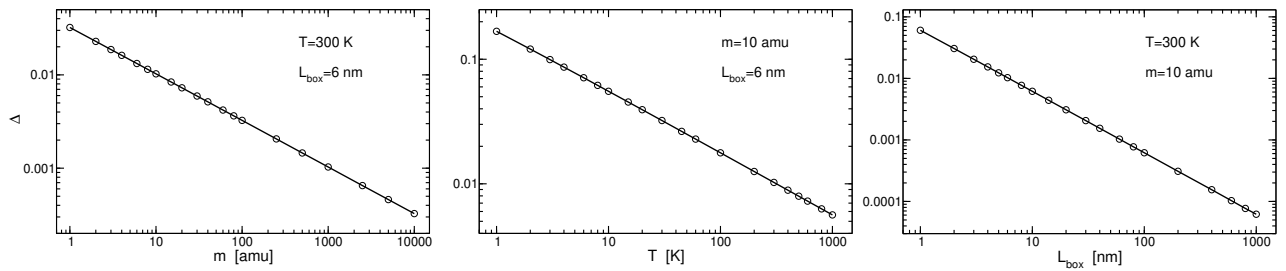


Figure SI-2.1: The relative error, Δ , defined in Eq. SI-2.6, of applying the 'classical' translation approximation to the ratio of the partition functions as a function of mass (left panel), temperature (middle panel), and box-length (right panel). The values of the parameters fixed in each plot correspond to the system defined in Section SI-1.

The smallest mass considered is 1 *amu* which corresponds to the lightest (i.e., a hydrogen) atom.

For temperature, the lowest value shown is 1 K . Although attainable, is unlikely to be of interest for most association/dimerization reactions of molecular systems and higher temperatures are more relevant. The smallest value considered for volume is that which corresponds to $L_{box} = 1 \text{ nm}$. Again, this smallest system is not likely to be applicable here because it can not satisfy the ideal gas behavior assumed in the derivation of K . Given the excluded volume of one atom and range of interaction between two atoms, a larger system is required. For the two-site monomer system described in the main text, with a LJ potential acting between particles (i.e., the dispersions decay as $1/r^6$), we found that a box length of 3 – 4 nm is probably the smallest for which ideal gas behavior can be observed. In any case in Fig. SI-2.1, the largest relative error observed is 17 % (middle panel at $T = 1 \text{ K}$) indicating the approximation in this case is not valid.

We now attempt to identify chemical systems for which the 'classical' translation approximation will exhibit the largest deviations. Very low temperatures are crucial, and systems operative under this condition are low molecular weight gases just above their boiling temperature. In Table SI-2.1 we list four gases (helium, hydrogen, neon, and nitrogen) having the lowest boiling points (4 – 77 K). We then consider these gases in a small box, that in our computational experience is already too small to support ideal behavior, and calculate the relative error Δ . What should be considered an acceptable error? Because ΔG^\ominus is related to K by a natural logarithm, a given error in the value of the latter translates to a much lower error of the former. We therefore propose, arbitrarily, relative errors lower than 0.05 to be acceptable and mark larger errors in table SI-2.1 by red color. For hydrogen gas, only at temperatures higher than $\sim 200 \text{ K}$ the 'classical' approximation can be applied, for helium, at temperatures higher than $\sim 100 \text{ K}$, whereas for neon and nitrogen, or for any other gas, at any temperature.

Table SI-2.1: The relative error, Δ , defined in Eq. SI-2.6, of applying the 'classical' translation approximation to the ratio of the partition functions for dimerization of hydrogen (H_2), Helium (He), Neon (Ne), and Nitrogen (N_2) gases confined to a cubic box with $L_{\text{box}} = 3.0 \text{ nm}$, at their corresponding boiling point T_b and at three higher temperatures. Discrepancies with relative magnitude larger than an arbitrary threshold of 5 % are marked in red.

gas	m [amu]	T_b [K]	$\Delta(T = T_b)$	$\Delta(T = 100 \text{ K})$	$\Delta(T = 200 \text{ K})$	$\Delta(T = 300 \text{ K})$
H_2	2.0	20.3	0.17	0.078	0.055	0.045
He	4.0	4.2	0.25	0.055	0.039	0.032
Ne	20.2	27.1	0.047	0.025	0.018	0.014
N_2	28.0	77.4	0.024	0.021	0.015	0.012

References

- [1] McQuarrie, D. A. *Statistical Thermodynamics*; University Science Books: Mill Valley, CA, 1973.
- [2] Zangi, R. Binding Reactions at Finite Systems, *Phys. Chem. Chem. Phys.* **2022**, *24*, 9921–9929.